

WHAT IS CLAIMED IS:

- 1 1. A method of manufacturing a family of antigenic peptides comprising:
2 locating a plurality of variable positions in a region of a pathogen protein;
3 choosing a peptide sequence of the pathogen protein including the plurality of
4 variable positions;
5 selecting one or more substitute amino acid residues for one of the variable positions
6 based on antigenic similarity to amino acid residues naturally occurring at the variable
7 position of the pathogen protein; and
8 preparing a family of antigenic peptides based on the peptide sequence and including
9 the substitute amino acid residues.
- 1 2. The method of claim 1, wherein selecting includes determining the antigenic
2 similarity using an antigenic similarity matrix.
- 1 3. The method of claim 1, further comprising assigning a frequency to each
2 substitute amino acid residue in the family of antigenic peptides.
- 1 4. The method of claim 3, wherein preparing includes weighting the substitute
2 amino acid residues in the family of antigenic peptides based on the assigned frequency.
- 1 5. The method of claim 3, wherein assigning further includes considering the
2 frequency with which the variations naturally occur.
- 1 6. The method of claim 1, wherein the pathogen protein includes a hypervariable
2 region.
- 1 7. The method of claim 1, wherein the pathogen protein is associated with a
2 virus.
- 1 8. The method of claim 7, wherein the virus is HIV.
- 1 9. The method of claim 1, wherein the pathogen protein is HIV gp120.

1 10. The method of claim 9, wherein the region is selected from the group
2 consisting of the V1, V2, V3, V4, and V5 regions.

1 11. The method of claim 7, wherein the virus is hepatitis B virus or hepatitis C
2 virus.

1 12. The method of claim 7, wherein the virus is an influenza virus.

1 13. The method of claim 7, wherein the virus is a dengue virus.

1 14. The method of claim 1, wherein the pathogen protein is associated with a
2 malaria pathogen.

1 15. The method of claim 1, wherein the pathogen protein is associated with a
2 tuberculosis pathogen.

1 16. The method of claim 1, further comprising determining an antigenic similarity
2 between a peptide of the family and a region of a human protein.

1 17. The method of claim 16, further comprising removing a peptide from the
2 family of antigenic peptides before preparing the family if the determined antigenic similarity
3 between the peptide of the family and the region of a human protein exceeds a predetermined
4 threshold.

1 18. The method of claim 1, wherein the family of antigenic peptides includes
2 members, such that the members taken together have antigenic similarity to each naturally
3 occurring sequence of the region of the pathogen protein.

1 19. The method of claim 1, wherein the family of antigenic peptides includes
2 members, such that the members taken together have antigenic similarity to a non-naturally
3 occurring sequence of the region of the pathogen protein.

1 20. The method of claim 1, further comprising identifying peptide sequences of
2 the family, the identified peptide sequences being representative of the sequence diversity of
3 the entire family.

1 21. The method of claim 20, wherein fewer than 500 sequences are identified.

1 22. The method of claim 20, wherein identifying includes calculating a distance
2 between peptide sequences of the family.

1 23. The method of claim 22, wherein calculating a distance includes using an
2 antigenic similarity matrix.

1 24. The method of claim 1, wherein preparing the family of antigenic peptides
2 includes chemical synthesis of the family of peptides.

1 25. The method of claim 24, wherein the chemical synthesis includes
2 combinatorial synthesis, whereby the peptides are formed as a mixture of different
3 sequences.

1 26. The method of claim 24, wherein the chemical synthesis includes parallel
2 synthesis, whereby each peptide is formed separately from other peptides.

1 27. The method of claim 26, further comprising mixing the separate peptides.

1 28. The method of claim 1, wherein preparing the family of antigenic peptides
2 includes expression of the family of peptides by a host organism.

1 29. A composition comprising a family of antigenic peptides having amino acid
2 sequences having antigenic similarity to amino acid sequences of a variable region of a
3 pathogen protein, wherein each antigenic peptide in the family has at least one amino acid
4 position that varies relative to other antigenic peptides in the family.

1 30. The composition of claim 29, wherein one amino acid residue occurs more
2 frequently than another in the position that varies.

1 31. The composition of claim 29, wherein the family includes greater than 150
2 mutually unique antigenic peptides.

1 32. The composition of claim 29, wherein the family includes greater than 1,000
2 mutually unique antigenic peptides.

1 33. The composition of claim 29, wherein the family includes fewer than 100,000
2 mutually unique antigenic peptides.

1 34. The composition of claim 29, wherein the family includes fewer than 50,000
2 mutually unique antigenic peptides.

1 35. The composition of claim 29, wherein the family includes between 1,000 and
2 50,000 mutually unique antigenic peptides.

1 36. The composition of claim 29, wherein the pathogen protein is HIV gp120.

1 37. The composition of claim 36, wherein the variable region is selected from the
2 group consisting of the V1 region, the V2 region, the V3 region, the V4 region and the V5
3 region.

1 38. The composition of claim 37, wherein the family of antigenic peptides
2 includes sequences having antigenic similarity to sequences from a subtype of HIV.

1 39. The composition of claim 38, wherein the subtype is selected from the group
2 consisting of subtype A, subtype B, subtype C, subtype D, subtype F, subtype G, a
3 recombinant subtype, a subtype of HIV group N, a subtype of HIV group O, and
4 combinations thereof.

1 40. The composition of claim 29, wherein at least two members of the family of
2 antigenic peptides are mixed together.

1 41. The composition of claim 29, wherein the family of antigenic peptides are
2 separated according to sequence.

1 42. The composition of claim 29, wherein the family includes a multiple antigenic
2 peptide.

1 43. A peptide library comprising a family of peptides including the fragment:
2 -N-N-T-R-X⁴-X⁵-X⁶-X⁷-X⁸-G-X⁹-G-X¹⁰-X¹¹-X¹²-X¹³-X¹⁴-T-G-X¹⁵-I-X¹⁶-G-X¹⁷-I-R-
3 wherein each X⁴-X¹⁷ is a fragment zero, one, two or three amino acid residues in length.

1 44. The peptide library of claim 43, wherein the family has antigenic similarity to
2 the V3 region of HIV gp120.

1 45. The peptide library of claim 44, wherein the family has antigenic similarity to
2 the V3 region of HIV gp120 of HIV subtypes A, B, C, D, F, and G.

1 46. The peptide library of claim 43, wherein the family of peptides have the
2 formula:

3 $X^1-C-X^2-R-P-X^3-N-N-T-R-X^4-X^5-X^6-X^7-X^8-G-X^9-G-X^{10}-X^{11}-X^{12}-X^{13}-X^{14}-T-G-X^{15}-I-X^{16}-G-X^{17}-I-R-$
4 $X^{18}-A-X^{19}-C-X^{20}$

5 wherein each X^1-X^{20} is a fragment zero, one, two or three amino acid residues in length.

1 47. The peptide library of claim 46, wherein for each peptide of the family, X^1
2 independently is N, V, T or E.

1 48. The peptide library of claim 46, wherein for each peptide of the family, X^2
2 independently is T or I.

1 49. The peptide library of claim 46, wherein for each peptide of the family, X^3
2 independently is N, S, G or Y.

1 50. The peptide library of claim 46, wherein for each peptide of the family, X^4
2 independently is K, Q or T.

1 51. The peptide library of claim 46, wherein for each peptide of the family, X^5
2 independently is S, G or R.

1 52. The peptide library of claim 46, wherein for each peptide of the family, X^6
2 independently is I, T or M.

1 53. The peptide library of claim 46, wherein for each peptide of the family, X^7
2 independently is H, R, P, S or N.

1 54. The peptide library of claim 46, wherein for each peptide of the family, X^8
2 independently is I or M.

1 55. The peptide library of claim 46, wherein for each peptide of the family, X⁹
2 independently is P or L.

1 56. The peptide library of claim 46, wherein for each peptide of the family, X¹⁰
2 independently is R or Q.

1 57. The peptide library of claim 46, wherein for each peptide of the family, X¹¹
2 independently is A, T or V.

1 58. The peptide library of claim 46, wherein for each peptide of the family, X¹²
2 independently is F or W.

1 59. The peptide library of claim 46, wherein for each peptide of the family, X¹³
2 independently is Y, F or H.

1 60. The peptide library of claim 46, wherein for each peptide of the family, X¹⁴
2 independently is A or T.

1 61. The peptide library of claim 46, wherein for each peptide of the family, X¹⁵
2 independently is D, Q, A or R.

1 62. The peptide library of claim 46, wherein for each peptide of the family, X¹⁶
2 independently is I or T.

1 63. The peptide library of claim 46, wherein for each peptide of the family, X¹⁷
2 independently is D or N.

1 64. The peptide library of claim 46, wherein for each peptide of the family, X¹⁸
2 independently is Q or K.

1 65. The peptide library of claim 46, wherein for each peptide of the family, X¹⁹
2 independently is H or Y.

1 66. The peptide library of claim 46, wherein for each peptide of the family, X²⁰
2 independently is N or T.

1 67. The peptide library of claim 46, wherein for each peptide of the family, X¹
 2 independently is N, V, T or E; X² independently is T or I; X³ independently is N, S, G or Y;
 3 X⁴ independently is K, Q or T; X⁵ independently is S, G or R; X⁶ independently is I, T or M;
 4 X⁷ independently is H, R, P, S or N; X⁸ independently is I or M; X⁹ independently is P or L;
 5 X¹⁰ independently is R or Q; X¹¹ independently is A, T or V; X¹² independently is F or W;
 6 X¹³ independently is Y, F or H; X¹⁴ independently is A or T; X¹⁵ independently is D, Q, A or
 7 R; X¹⁶ independently is I or T; X¹⁷ independently is D or N; X¹⁸ independently is Q or K; X¹⁹
 8 independently is H or Y; and X²⁰ independently is N or T.

1 68. The peptide library of claim 46, wherein at least two members of the family of
 2 peptides are mixed together.

1 69. The peptide library of claim 46, wherein the family of peptides are separated
 2 according to sequence.

1 70. The peptide library of claim 46, wherein the family includes greater than 150
 2 mutually unique peptide sequences.

1 71. The peptide library of claim 46, wherein the family includes fewer than
 2 100,000 mutually unique peptide sequences.

1 72. The peptide library of claim 46, wherein the family includes fewer than 500
 2 mutually unique peptide sequences, the sequences being representative of the entire sequence
 3 diversity available.

1 73. A peptide library comprising a family of peptides including the fragment:
 2 -N-N-X⁴-R-X⁵-X⁶-V-X⁷-I-G-P-G-X⁸-X⁹-F-X¹⁰-X¹¹-X¹²-X¹³-X¹⁴-I-X¹⁵-G-X¹⁶-I-R-
 3 wherein each X⁴-X¹⁶ is a fragment zero, one, two or three amino acid residues in length.

1 74. The peptide library of claim 73, wherein the family has antigenic similarity to
 2 the V3 region of HIV gp120.

1 75. The peptide library of claim 74, wherein the family has antigenic similarity to
 2 the V3 region of HIV gp120 of HIV subtype A.

1 76. The peptide library of claim 73, wherein the family of peptides have the
2 formula:

3 $X^1-C-X^2-R-P-X^3-N-N-X^4-R-X^5-X^6-V-X^7-I-G-P-G-X^8-X^9-F-X^{10}-X^{11}-X^{12}-X^{13}-X^{14}-I-X^{15}-G-X^{16}-I-R-$
4 $X^{17}-A-X^{18}-C-X^{19}$

5 wherein each X^1-X^{19} is a fragment zero, one, two or three amino acid residues in length.

1 77. The peptide library of claim 76, wherein for each peptide of the family, X^1
2 independently is N, T, or D.

1 78. The peptide library of claim 76, wherein for each peptide of the family, X^2
2 independently is T or I.

1 79. The peptide library of claim 76, wherein for each peptide of the family, X^3
2 independently is N, G, or S.

1 80. The peptide library of claim 76, wherein for each peptide of the family, X^4
2 independently is T or K.

1 81. The peptide library of claim 76, wherein for each peptide of the family, X^5
2 independently is K, T, or Q.

1 82. The peptide library of claim 76, wherein for each peptide of the family, X^6
2 independently is S, G, R or N.

1 83. The peptide library of claim 76, wherein for each peptide of the family, X^7
2 independently is R or H.

1 84. The peptide library of claim 76, wherein for each peptide of the family, X^8
2 independently is Q or R.

1 85. The peptide library of claim 76, wherein for each peptide of the family, X^9
2 independently is A, T, or V.

1 86. The peptide library of claim 76, wherein for each peptide of the family, X^{10}
2 independently is Y, F, or H.

1 87. The peptide library of claim 76, wherein for each peptide of the family, X¹¹
2 independently is A or T.

1 88. The peptide library of claim 76, wherein for each peptide of the family, X¹²
2 independently is T, R, A, or no residue.

1 89. The peptide library of claim 76, wherein for each peptide of the family, X¹³
2 independently is G, D, or no residue.

1 90. The peptide library of claim 76, wherein for each peptide of the family, X¹⁴
2 independently is D, A, N, R, or V.

1 91. The peptide library of claim 76, wherein for each peptide of the family, X¹⁵
2 independently is I or T.

1 92. The peptide library of claim 76, wherein for each peptide of the family, X¹⁶
2 independently is D or N.

1 93. The peptide library of claim 76, wherein for each peptide of the family, X¹⁷
2 independently is Q, K, or E.

1 94. The peptide library of claim 76, wherein for each peptide of the family, X¹⁸
2 independently is H or Y.

1 95. The peptide library of claim 76, wherein for each peptide of the family, X¹⁹
2 independently is N, T, or E.

1 96. The peptide library of claim 76, wherein for each peptide of the family X¹
2 independently is N, T, or D; X² independently is T or I; X³ independently is N, G, or S; X⁴
3 independently is T or K; X⁵ independently is K, T, or Q; X⁶ independently is S, G, R or N;
4 X⁷ independently is R or H; X⁸ independently is Q or R; X⁹ independently is A, T, or V; X¹⁰
5 independently is Y, F, or H; X¹¹ independently is A or T; X¹² independently is T, R, A, or no
6 residue; X¹³ independently is G, D, or no residue; X¹⁴ independently is D, A, N, R, or V; X¹⁵
7 independently is I or T; X¹⁶ independently is D or N; X¹⁷ independently is Q, K, or E; X¹⁸
8 independently is H or Y; and X¹⁹ independently is N, T, or E.

1 97. The peptide library of claim 76, wherein at least two members of the family of
2 peptides are mixed together.

1 98. The peptide library of claim 76, wherein the family of peptides are separated
2 according to sequence.

1 99. The peptide library of claim 76, wherein the family includes greater than 150
2 mutually unique peptide sequences.

1 100. The peptide library of claim 76, wherein the family includes fewer than
2 100,000 mutually unique peptide sequences.

1 101. The peptide library of claim 76, wherein the family includes fewer than 500
2 mutually unique peptide sequences, the sequences being representative of the entire sequence
3 diversity available.

1 102. A peptide library comprising a family of peptides including the fragment:
2 -N-N-T-R-K-X⁴-I-X⁵-X⁶-G-X⁷-G-X⁸-X⁹-X¹⁰-X¹¹-X¹²-T-X¹³-X¹⁴-I-X¹⁵-G-X¹⁶-I-R-
3 wherein each X⁴-X¹⁶ is a fragment zero, one, two or three amino acid residues in length.

1 103. The peptide library of claim 102, wherein the family has antigenic similarity
2 to the V3 region of HIV gp120.

1 104. The peptide library of claim 103, wherein the family has antigenic similarity
2 to the V3 region of HIV gp120 of HIV subtype B.

1 105. The peptide library of claim 102, wherein the family of peptides have the
2 formula:

3 X¹-C-X²-R-P-X³-N-N-T-R-K-X⁴-I-X⁵-X⁶-G-X⁷-G-X⁸-X⁹-X¹⁰-X¹¹-X¹²-T-X¹³-X¹⁴-I-X¹⁵-G-X¹⁶-I-R-
4 X¹⁷-A-X¹⁸-C-X¹⁹
5 wherein each X¹-X¹⁹ is a fragment zero, one, two or three amino acid residues in length.

1 106. The peptide library of claim 105, wherein for each peptide of the family, X¹
2 independently is N, T, or H.

1 107. The peptide library of claim 105, wherein for each peptide of the family, X²
2 independently is T or I.

1 108. The peptide library of claim 105, wherein for each peptide of the family, X³
2 independently is N, S, or G.

1 109. The peptide library of claim 105, wherein for each peptide of the family, X⁴
2 independently is S, G, or R.

1 110. The peptide library of claim 105, wherein for each peptide of the family, X⁵
2 independently is H, P, N, T or Y.

1 111. The peptide library of claim 105, wherein for each peptide of the family, X⁶
2 independently is I or M.

1 112. The peptide library of claim 105, wherein for each peptide of the family, X⁷
2 independently is P, L, or W.

1 113. The peptide library of claim 105, wherein for each peptide of the family, X⁸
2 independently is R, Q, G or S.

1 114. The peptide library of claim 105, wherein for each peptide of the family, X⁹
2 independently is A, V or T.

1 115. The peptide library of claim 105, wherein for each peptide of the family, X¹⁰
2 independently is F, W, or V.

1 116. The peptide library of claim 105, wherein for each peptide of the family, X¹¹
2 independently is Y, F or H.

1 117. The peptide library of claim 105, wherein for each peptide of the family, X¹²
2 independently is T or A.

1 118. The peptide library of claim 105, wherein for each peptide of the family, X¹³
2 independently is G, E or R.

1 119. The peptide library of claim 105, wherein for each peptide of the family, X¹⁴
2 independently is E, Q, R or G.

1 120. The peptide library of claim 105, wherein for each peptide of the family, X¹⁵
2 independently is I or T.

1 121. The peptide library of claim 105, wherein for each peptide of the family, X¹⁶
2 independently is D or N.

1 122. The peptide library of claim 105, wherein for each peptide of the family, X¹⁷
2 independently is Q or K.

1 123. The peptide library of claim 105, wherein for each peptide of the family, X¹⁸
2 independently is H or Y.

1 124. The peptide library of claim 105, wherein for each peptide of the family, X¹⁹
2 independently is N or T.

1 125. The peptide library of claim 105, wherein for each peptide of the family, X¹
2 independently is N, T or H; X² independently is T or I; X³ independently is N, S, or G; X⁴
3 independently is S, G, or R; X⁵ independently is H, P, N, T or Y; X⁶ independently is I or M;
4 X⁷ independently is P, L, or W; X⁸ independently is R, Q, G or S; X⁹ independently is A, V
5 or T; X¹⁰ independently is F, W, or V; X¹¹ independently is Y, F or H; X¹² independently is T
6 or A; X¹³ independently is G, E or R; X¹⁴ independently is E, Q, R or G; X¹⁵ independently is
7 I or T; X¹⁶ independently is D or N; X¹⁷ independently is Q or K; X¹⁸ independently is H or
8 Y; and X¹⁹ independently is N or T.

1 126. The peptide library of claim 105, wherein at least two members of the family
2 of peptides are mixed together.

1 127. The peptide library of claim 105, wherein the family of peptides are separated
2 according to sequence.

1 128. The peptide library of claim 105, wherein the family includes greater than 150
2 mutually unique peptide sequences.

1 129. The peptide library of claim 105, wherein the family includes fewer than
2 100,000 mutually unique peptide sequences.

1 130. The peptide library of claim 105, wherein the family includes fewer than 500
2 mutually unique peptide sequences, the sequences being representative of the entire sequence
3 diversity available.

1 131. A peptide library comprising a family of peptides including the fragment:
2 -N-N-T-R-X⁴-X⁵-X⁶-X⁷-I-G-P-G-Q-X⁸-F-X⁹-X¹⁰-T-X¹¹-X¹²-I-X¹³-G-X¹⁴-I-R-
3 wherein each X⁴-X¹⁴ is a fragment zero, one, two or three amino acid residues in length.

1 132. The peptide library of claim 131, wherein the family has antigenic similarity
2 to the V3 region of HIV gp120.

1 133. The peptide library of claim 132, wherein the family has antigenic similarity
2 to the V3 region of HIV gp120 of HIV subtype C.

1 134. The peptide library of claim 131, wherein the family of peptides have the
2 formula:
3 X¹-C-X²-R-P-X³-N-N-T-R-X⁴-X⁵-X⁶-X⁷-I-G-P-G-Q-X⁸-F-X⁹-X¹⁰-T-X¹¹-X¹²-I-X¹³-G-X¹⁴-I-R-X¹⁵-A-
4 X¹⁶-C-X¹⁷
5 wherein each X¹-X¹⁷ is a fragment zero, one, two or three amino acid residues in length.

1 135. The peptide library of claim 134, wherein for each peptide of the family, X¹
2 independently is V, N, E, T, M or R.

1 136. The peptide library of claim 134, wherein for each peptide of the family, X²
2 independently is T, I, or A.

1 137. The peptide library of claim 134, wherein for each peptide of the family, X³
2 independently is N, G, S, or H.

1 138. The peptide library of claim 134, wherein for each peptide of the family, X⁴
2 independently is K, E or Q.

1 139. The peptide library of claim 134, wherein for each peptide of the family, X⁵
2 independently is S or G.

1 140. The peptide library of claim 134, wherein for each peptide of the family, X⁶
2 independently is I or M.

1 141. The peptide library of claim 134, wherein for each peptide of the family, X⁷
2 independently is R or G.

1 142. The peptide library of claim 134, wherein for each peptide of the family, X⁸
2 independently is T, A, or V.

1 143. The peptide library of claim 134, wherein for each peptide of the family, X⁹
2 independently is Y or F.

1 144. The peptide library of claim 134, wherein for each peptide of the family, X¹⁰
2 independently is A or T.

1 145. The peptide library of claim 134, wherein for each peptide of the family, X¹¹
2 independently is G, N, D, K or T.

1 146. The peptide library of claim 134, wherein for each peptide of the family, X¹²
2 independently is D, G, N, S, or I.

1 147. The peptide library of claim 134, wherein for each peptide of the family, X¹³
2 independently is I or T.

1 148. The peptide library of claim 134, wherein for each peptide of the family, X¹⁴
2 independently is D or N.

1 149. The peptide library of claim 134, wherein for each peptide of the family, X¹⁵
2 independently is Q, K, E, L or H.

1 150. The peptide library of claim 134, wherein for each peptide of the family, X¹⁶
2 independently is H, Y, or N.

1 151. The peptide library of claim 134, wherein for each peptide of the family, X¹⁷
2 independently is N, T, I, D or H.

1 152. The peptide library of claim 134, wherein for each peptide of the family, X¹
2 independently is V, N, E, T, M or R; X² independently is T, I, or A; X³ independently is N,
3 G, S, or H; X⁴ independently is K, E or Q; X⁵ independently is S or G; X⁶ independently is I
4 or M; X⁷ independently is R or G; X⁸ independently is T, A, or V; X⁹ independently is Y or
5 F; X¹⁰ independently is A or T; X¹¹ independently is G, N, D, K or T; X¹² independently is
6 D, G, N, S, or I; X¹³ independently is I or T; X¹⁴ independently is D or N; X¹⁵ independently
7 is Q, K, E, L or H; X¹⁶ independently is H, Y, or N; and X¹⁷ independently is N, T, I, D or H.

1 153. The peptide library of claim 134, wherein at least two members of the family
2 of peptides are mixed together.

1 154. The peptide library of claim 134, wherein the family of peptides are separated
2 according to sequence.

1 155. The peptide library of claim 134, wherein the family includes greater than 150
2 mutually unique peptide sequences.

1 156. The peptide library of claim 134, wherein the family includes fewer than
2 100,000 mutually unique peptide sequences.

1 157. The peptide library of claim 134, wherein the family includes fewer than 500
2 mutually unique peptide sequences, the sequences being representative of the entire sequence
3 diversity available.

1 158. A peptide library comprising a family of peptides including the fragment:
2 -R-P-X³-X⁴-X⁵-X⁶-R-X⁷-X⁸-X⁹-X¹⁰-I-G-X¹¹-G-X¹²-X¹³-X¹⁴-X¹⁵-X¹⁶-T-X¹⁷-X¹⁸-G-X¹⁹-I-
3 wherein each X³-X¹⁹ is a fragment zero, one, two or three amino acid residues in length.

1 159. The peptide library of claim 158, wherein the family has antigenic similarity
2 to the V3 region of HIV gp120.

1 160. The peptide library of claim 159, wherein the family has antigenic similarity
2 to the V3 region of HIV gp120 of subtype D.

1 161. The peptide library of claim 158, wherein the family of peptides have the
2 formula:

3 $X^1-C-X^2-R-P-X^3-X^4-X^5-X^6-R-X^7-X^8-X^9-X^{10}-I-G-X^{11}-G-X^{12}-X^{13}-X^{14}-X^{15}-X^{16}-T-X^{17}-X^{18}-G-X^{19}-I-X^{20}-$
4 $X^{21}-A-X^{22}-C-X^{23}$

5 wherein each X^1-X^{23} is a fragment zero, one, two or three amino acid residues in length.

1 162. The peptide library of claim 161, wherein for each peptide of the family, X^1
2 independently is N or T.

1 163. The peptide library of claim 161, wherein for each peptide of the family, X^2
2 independently is T or I.

1 164. The peptide library of claim 161, wherein for each peptide of the family, X^3
2 independently is Y or N.

1 165. The peptide library of claim 161, wherein for each peptide of the family, X^4
2 independently is N or K.

1 166. The peptide library of claim 161, wherein for each peptide of the family, X^5
2 independently is N or K.

1 167. The peptide library of claim 161, wherein for each peptide of the family, X^6
2 independently is T, I or K.

1 168. The peptide library of claim 161, wherein for each peptide of the family, X^7
2 independently is Q or R.

1 169. The peptide library of claim 161, wherein for each peptide of the family, X^8
2 independently is S, R or G.

1 170. The peptide library of claim 161, wherein for each peptide of the family, X^9
2 independently is T or I.

1 171. The peptide library of claim 161, wherein for each peptide of the family, X¹⁰
2 independently is H, P, S or R.

1 172. The peptide library of claim 161, wherein for each peptide of the family, X¹¹
2 independently is P, L, S or Q.

1 173. The peptide library of claim 161, wherein for each peptide of the family, X¹²
2 independently is Q or R.

1 174. The peptide library of claim 161, wherein for each peptide of the family, X¹³
2 independently is A or T.

1 175. The peptide library of claim 161, wherein for each peptide of the family, X¹⁴
2 independently is L or Y.

1 176. The peptide library of claim 161, wherein for each peptide of the family, X¹⁵
2 independently is Y or F.

1 177. The peptide library of claim 161, wherein for each peptide of the family, X¹⁶
2 independently is T or A.

1 178. The peptide library of claim 161, wherein for each peptide of the family, X¹⁷
2 independently is I or K.

1 179. The peptide library of claim 161, wherein for each peptide of the family, X¹⁸
2 independently is I, T or K.

1 180. The peptide library of claim 161, wherein for each peptide of the family, X¹⁹
2 independently is D or N.

1 181. The peptide library of claim 161, wherein for each peptide of the family, X²⁰
2 independently is R or G.

1 182. The peptide library of claim 161, wherein for each peptide of the family, X²¹
2 independently is Q or K.

1 183. The peptide library of claim 161, wherein for each peptide of the family, X²²
2 independently is H or Y.

1 184. The peptide library of claim 161, wherein for each peptide of the family, X²³
2 independently is N, T or K.

1 185. The peptide library of claim 161, wherein for each peptide of the family, X¹
2 independently is N or T; X² independently is T or I; X³ independently is Y or N; X⁴
3 independently is N or K; X⁵ independently is N or K; X⁶ independently is T, I or K; X⁷
4 independently is Q or R; X⁸ independently is S, R or G; X⁹ independently is T or I; X¹⁰
5 independently is H, P, S or R; X¹¹ independently is P, L, S or Q; X¹² independently is Q or R;
6 X¹³ independently is A or T; X¹⁴ independently is L or Y; X¹⁵ independently is Y or F; X¹⁶
7 independently is T or A; X¹⁷ independently is I or K; X¹⁸ independently is I, T or K; X¹⁹
8 independently is D or N; X²⁰ independently is R or G; X²¹ independently is Q or K; X²²
9 independently is H or Y; and X²³ independently is N, T or K.

1 186. The peptide library of claim 161, wherein at least two members of the family
2 of peptides are mixed together.

1 187. The peptide library of claim 161, wherein the family of peptides are separated
2 according to sequence.

1 188. The peptide library of claim 161, wherein the family includes greater than 150
2 mutually unique peptide sequences.

1 189. The peptide library of claim 161, wherein the family includes fewer than
2 100,000 mutually unique peptide sequences.

1 190. The peptide library of claim 161, wherein the family includes fewer than 500
2 mutually unique peptide sequences, the sequences being representative of the entire sequence
3 diversity available.

1 191. A peptide library comprising a family of peptides including the fragment:
2 -N-N-X³-R-K-X⁴-I-X⁵-L-G-P-G-X⁶-X⁷-X⁸-X⁹-X¹⁰-X¹¹-X¹²-X¹³-I-X¹⁴-G-X¹⁵-I-R-
3 wherein each X³-X¹⁵ is a fragment zero, one, two or three amino acid residues in length.

1 192. The peptide library of claim 191, wherein the family has antigenic similarity
2 the V3 region of HIV gp120.

1 193. The peptide library of claim 192, wherein the family has antigenic similarity
2 the V3 region of HIV gp120 of HIV subtype F.

1 194. The peptide library of claim 191, wherein the family of peptides have the
2 formula:

3 $X^1\text{-C-T-R-P-X}^2\text{-N-N-X}^3\text{-R-K-X}^4\text{-I-X}^5\text{-L-G-P-G-X}^6\text{-X}^7\text{-X}^8\text{-X}^9\text{-X}^{10}\text{-X}^{11}\text{-X}^{12}\text{-X}^{13}\text{-I-X}^{14}\text{-G-X}^{15}\text{-I-R-X}^{16}\text{-}$
4 $\text{A-X}^{17}\text{-C-X}^{18}$

5 wherein each $X^1\text{-X}^{18}$ is a fragment zero, one, two or three amino acid residues in length.

1 195. The peptide library of claim 194, wherein for each peptide of the family, X^1
2 independently is N or T.

1 196. The peptide library of claim 194, wherein for each peptide of the family, X^2
2 independently is N or S.

1 197. The peptide library of claim 194, wherein for each peptide of the family, X^3
2 independently is T or I.

1 198. The peptide library of claim 194, wherein for each peptide of the family, X^4
2 independently is S, G, or R.

1 199. The peptide library of claim 194, wherein for each peptide of the family, X^5
2 independently is H, R, Q, P, Y, or S.

1 200. The peptide library of claim 194, wherein for each peptide of the family, X^6
2 independently is Q, R or H.

1 201. The peptide library of claim 194, wherein for each peptide of the family, X^7
2 independently is A, V or T.

1 202. The peptide library of claim 194, wherein for each peptide of the family, X^8
2 independently is F or I.

1 203. The peptide library of claim 194, wherein for each peptide of the family, X⁹
2 independently is Y, H or F.

1 204. The peptide library of claim 194, wherein for each peptide of the family, X¹⁰
2 independently is A or T.

1 205. The peptide library of claim 194, wherein for each peptide of the family, X¹¹
2 independently is T or A.

1 206. The peptide library of claim 194, wherein for each peptide of the family, X¹²
2 independently is G, D or S.

1 207. The peptide library of claim 194, wherein for each peptide of the family, X¹³
2 independently is D, A, or N.

1 208. The peptide library of claim 194, wherein for each peptide of the family, X¹⁴
2 independently is I or T.

1 209. The peptide library of claim 194, wherein for each peptide of the family, X¹⁵
2 independently is D or N.

1 210. The peptide library of claim 194, wherein for each peptide of the family, X¹⁶
2 independently is K or Q.

1 211. The peptide library of claim 194, wherein for each peptide of the family, X¹⁷
2 independently is H or Y.

1 212. The peptide library of claim 194, wherein for each peptide of the family, X¹⁸
2 independently is N, I, D, or K.

1 213. The peptide library of claim 194, wherein for each peptide of the family, X¹
2 independently is N or T; X² independently is N or S; X³ independently is T or I; X⁴
3 independently is S, G, or R; X⁵ independently is H, R, Q, P, Y, or S; X⁶ independently is Q,
4 R or H; X⁷ independently is A, V or T; X⁸ independently is F or I; X⁹ independently is Y, H
5 or F; X¹⁰ independently is A or T; X¹¹ independently is T or A; X¹² independently is G, D or

6 S; X¹³ independently is D, A, or N; X¹⁴ independently is I or T; X¹⁵ independently is D or N;
7 X¹⁶ independently is K or Q; X¹⁷ independently is H or Y; and X¹⁸ independently is N, I, D,
8 K or X.

1 214. The peptide library of claim 194, wherein at least two members of the family
2 of peptides are mixed together.

1 215. The peptide library of claim 194, wherein the family of peptides are separated
2 according to sequence.

1 216. The peptide library of claim 194, wherein the family includes greater than 150
2 mutually unique peptide sequences.

1 217. The peptide library of claim 194, wherein the family includes fewer than
2 100,000 mutually unique peptide sequences.

1 218. The peptide library of claim 194, wherein the family includes fewer than 500
2 mutually unique peptide sequences, the sequences being representative of the entire sequence
3 diversity available.

1 219. A peptide library comprising a family of peptides including the fragment:
2 -N-N-T-R-K-S-X⁴-X⁵-X⁶-G-X⁷-X⁸-X⁹-X¹⁰-X¹¹-X¹²-X¹³-T-X¹⁴-I-X¹⁵-G-X¹⁶-I-R-
3 wherein each X⁴-X¹⁶ is a fragment zero, one, two or three amino acid residues in length.

1 220. The peptide library of claim 219, wherein the family has antigenic similarity
2 to the V3 region of HIV gp120.

1 221. The peptide library of claim 220, wherein the family has antigenic similarity
2 to the V3 region of HIV gp120 of HIV subtype G.

1 222. The peptide library of claim 219, wherein the family of peptides have the
2 formula:

3 X¹-C-X²-R-P-X³-N-N-T-R-K-S-X⁴-X⁵-X⁶-G-X⁷-X⁸-X⁹-X¹⁰-X¹¹-X¹²-X¹³-T-X¹⁴-I-X¹⁵-G-X¹⁶-I-R-X¹⁷-
4 A-X¹⁸-C-X¹⁹

5 wherein each X¹-X¹⁹ is a fragment zero, one, two or three amino acid residues in length.

1 223. The peptide library of claim 222, wherein for each peptide of the family, X¹
2 independently is N, I, T, M or R.

1 224. The peptide library of claim 222, wherein for each peptide of the family, X²
2 independently is T or I.

1 225. The peptide library of claim 222, wherein for each peptide of the family, X³
2 independently is N or S.

1 226. The peptide library of claim 222, wherein for each peptide of the family, X⁴
2 independently is I or K.

1 227. The peptide library of claim 222, wherein for each peptide of the family, X⁵
2 independently is T, R, H, P or N.

1 228. The peptide library of claim 222, wherein for each peptide of the family, X⁶
2 independently is F or L.

1 229. The peptide library of claim 222, wherein for each peptide of the family, X⁷
2 independently is P, T, or L.

1 230. The peptide library of claim 222, wherein for each peptide of the family, X⁸
2 independently is G or R.

1 231. The peptide library of claim 222, wherein for each peptide of the family, X⁹
2 independently is Q, R, H or G.

1 232. The peptide library of claim 222, wherein for each peptide of the family, X¹⁰
2 independently is A, T or V.

1 233. The peptide library of claim 222, wherein for each peptide of the family, X¹¹
2 independently is F, I or S.

1 234. The peptide library of claim 222, wherein for each peptide of the family, X¹²
2 independently is Y or H.

1 235. The peptide library of claim 222, wherein for each peptide of the family, X¹³
2 independently is A or T.

1 236. The peptide library of claim 222, wherein for each peptide of the family, X¹⁴
2 independently is G, D or S.

1 237. The peptide library of claim 222, wherein for each peptide of the family, X¹⁵
2 independently is I or T.

1 238. The peptide library of claim 222, wherein for each peptide of the family, X¹⁶
2 independently is D or N.

1 239. The peptide library of claim 222, wherein for each peptide of the family, X¹⁷
2 independently is Q, L, K or P.

1 240. The peptide library of claim 222, wherein for each peptide of the family, X¹⁸
2 independently is H, Y or F.

1 241. The peptide library of claim 222, wherein for each peptide of the family, X¹⁹
2 independently is N, K or T.

1 242. The peptide library of claim 222, wherein for each peptide of the family, X¹
2 independently is N, I, T, M or R; X² independently is T or I; X³ independently is N or S; X⁴
3 independently is I or K; X⁵ independently is T, R, H, P or N; X⁶ independently is F or I; X⁷
4 independently is P, T, or L; X⁸ independently is G or R; X⁹ independently is Q, R, H or G;
5 X¹⁰ independently is A, T or V; X¹¹ independently is F, I or S; X¹² independently is Y or H;
6 X¹³ independently is A or T; X¹⁴ independently is G, D or S; X¹⁵ independently is I or T; X¹⁶
7 independently is D or N; X¹⁷ independently is Q, L, K or P; X¹⁸ independently is H, Y or F;
8 X¹⁹ independently is N, K or T.

1 243. The peptide library of claim 222, wherein at least two members of the family
2 of peptides are mixed together.

1 244. The peptide library of claim 222, wherein the family of peptides are separated
2 according to sequence.

1 245. The peptide library of claim 222, wherein the family includes greater than 150
2 mutually unique peptide sequences.

1 246. The peptide library of claim 222, wherein the family includes fewer than
2 100,000 mutually unique peptide sequences.

1 247. The peptide library of claim 222, wherein the family includes fewer than 500
2 mutually unique peptide sequences, the sequences being representative of the entire sequence
3 diversity available.

1 248. A method of eliciting an immune response in a subject comprising
2 administering to the subject a composition including a family of antigenic peptides having
3 amino acid sequences having antigenic similarity to amino acid sequences of a variable
4 region of a pathogen protein of a pathogen.

1 249. The method of claim 248, wherein the antigenic similarity is determined using
2 an antigenic similarity matrix.

1 250. The method of claim 248, wherein the composition is administered to a
2 subject prior to infection by the pathogen.

1 251. The method of claim 250, wherein the family of antigenic peptides has amino
2 acid sequences having antigenic similarity to amino acid sequences of a subtype of the
3 pathogen protein.

1 252. The method of claim 250, wherein the family of antigenic peptides has amino
2 acid sequences having antigenic similarity to amino acid sequences from more than one
3 subtype of the pathogen protein.

1 253. The method of claim 248, wherein the composition is administered to a
2 subject infected by the pathogen.

1 254. The method of claim 253, wherein the subject is infected by a subtype of the
2 pathogen.

1 255. The method of claim 254, wherein the family of antigenic peptides has amino
2 acid sequences having antigenic similarity to amino acid sequences from the subtype by
3 which the subject is infected.

1 256. A method of designing a family of peptide sequences, comprising:
2 locating a plurality of variable positions in a region of a pathogen protein;
3 choosing a peptide sequence of the pathogen protein including the variable positions;
4 and
5 selecting one or more substitute amino acid residues for one of the variable positions
6 of the peptide based on antigenic similarity to amino acid residues naturally occurring at the
7 variable position of the pathogen protein, thereby forming a family of peptide sequences.

1 257. The method of claim 256, wherein selecting includes determining the
2 antigenic similarity using an antigenic similarity matrix.

1 258. The method of claim 256, further comprising assigning a frequency to each
2 substitute amino acid residue in the family of antigenic peptide sequences.

1 259. The method of claim 258, wherein assigning further includes considering the
2 frequency with which the variations naturally occur.

1 260. The method of claim 256, further comprising determining an antigenic
2 similarity between a peptide of the family and a region of a human protein.

1 261. The method of claim 260, further comprising removing a peptide sequence
2 from the family of peptide sequences if the calculated antigenic similarity between the
3 peptide of the family and the region of a human protein exceeds a predetermined threshold.

1 262. The method of claim 256, wherein the family of peptide sequences includes
2 members, such that the members taken together have antigenic similarity to each naturally
3 occurring sequence of the region of the pathogen protein.

1 263. The method of claim 256, wherein the family of antigenic peptides includes
2 members, such that the members taken together have antigenic similarity to a non-naturally
3 occurring sequence of the region of the pathogen protein.

1 264. The method of claim 256, further comprising identifying peptide sequences of
2 the family, the identified peptide sequences being representative of the sequence diversity of
3 the entire family.

1 265. The method of claim 264, wherein fewer than 500 sequences are identified.

1 266. The method of claim 264, wherein identifying includes calculating a distance
2 between peptide sequences of the family.

1 267. The method of claim 266, wherein calculating a distance includes using an
2 antigenic similarity matrix.

1 268. The method of claim 256, further comprising identifying peptide sequences of
2 the family having antigenic similarity to a subtype of the pathogen protein.

1 269. The method of claim 256, wherein the pathogen protein includes a
2 hypervariable region.

1 270. The method of claim 256, wherein the pathogen is a virus.

1 271. The method of claim 270, wherein the virus is HIV.

1 272. The method of claim 271, wherein the pathogen protein is gp120.

1 273. The method of claim 272, wherein the region is selected from the group
2 consisting of the V1 region, the V2 region, the V3 region, the V4 region, and the V5 region.

1 274. A method of diagnosing infection comprising contacting a sample with a
2 family of peptides having amino acid sequences having antigenic similarity to amino acid
3 sequences of a variable region of a pathogen protein, wherein each peptide in the family has
4 at least one amino acid position that varies relative to other peptides in the family.

1 275. The method of claim 274, wherein antigenic similarity is determined using an
2 antigenic similarity matrix.

1 276. The method of claim 274, wherein the family of peptides is antigenically
2 similar to a pathogen subtype.

1 277. The method of claim 274, wherein the family of peptides is antigenically
2 similar to more than one pathogen subtype.

1 278. The method of claim 274, further comprising immobilizing the family of
2 peptides on a substrate before contacting.

1 279. The method of claim 274, further comprising determining if the sample
2 includes antibodies that bind specifically to the family of peptides.